



Let My People Go: Ethnic In-Group Bias in Judicial Decisions—Evidence from a Randomized Natural Experiment

*Oren Gazal-Ayal and Raanan Sulitzeanu-Kenan**

Does ethnic identity affect judicial decisions? We provide new evidence on ethnic biases in judicial behavior by examining the decisions of Arab and Jewish judges in first bail hearings of Arab and Jewish suspects in Israeli courts. Our setting avoids the potential bias from unobservable case characteristics by exploiting the random assignment of judges to cases during weekends and by focusing on the *difference* in ethnic disparity between Arab and Jewish judges. The study concentrates on the early-stage decisions in the judicial criminal process, controlling for the state's position and excluding agreements, thereby allowing us to distinguish judicial bias from other sources of ethnic disparities. We find systematic evidence of in-group (same ethnic group) bias in detention decisions. However, in cases where the decision is to detain, no ethnic bias was found in the length of the detention. Possible interpretations and implications of these findings are discussed.

I. INTRODUCTION

Most judicial decisions should not be affected by the race or ethnicity of the parties, and all fair judicial decisions should be oblivious to the judges' identity. Still, many studies point to the possible effect defendants' race and ethnicity have on judicial decisions, given racial and ethnic disparities in various punitive treatments. Yet inferring judicial bias from racial and ethnic disparities has proven difficult. The question remains whether these differences reflect racial differences in criminal behavior or the outcome of ethnic bias? More formally,

*Address correspondence to Raanan Sulitzeanu-Kenan, the Hebrew University of Jerusalem, Political Science Department, and Federmann School of Public Policy, Mount Scopus 91905, Israel; email: raulke@huji.ac.il. Gazal-Ayal is Senior Lecturer, University of Haifa, Faculty of Law; Sulitzeanu-Kenan is Lecturer, the Hebrew University of Jerusalem, Political Science Department, and Federmann School of Public Policy.

Earlier versions of this article were presented at faculty seminars of the Hebrew University Law Faculty, and Economics Department, Bar Ilan University, Faculty of Law, the Israeli Association of Law and Society 2008 Annual Conference, and at the conference "Empirical Studies of Criminal and Civil Procedures" at the Hebrew University of Jerusalem, May 2008. We are grateful for the suggestions and comments made there, and for comments by Adi Ayal, Momi Dahan, Ted Eisenberg, Gideon Fishman, Micha Mandel, Jacob Nussim, Danny Pfeffermann, Jeff Rachlinski, Doron Teichman, Yoram Shachar, Ron Shapira, Tamir Sheaffer, Keren Weinshall-Margel, David Weisburd, Gadi Wolfsfeld, and Asaf Zussman. Funding for this study was provided by the Minerva Center for Human Rights at the Hebrew University of Jerusalem. Additionally, we thank Roni Factor for his advice and assistance with the data analysis, Gal Einav and Attalla Shubash for collecting the data, and Yovav Eshet for research assistance. Any faults that remain are, of course, our own.

a null hypothesis of “zero disparity in judicial outcomes” can be rejected only if *all* the relevant variables are controlled for. Indeed, differential control of contextual variables in various studies has produced conflicting results.

Moreover, most of these studies concentrated on the effect of race or ethnicity on sentencing. Yet sentences are affected by many procedural decisions that precede this final stage of the trial. Among others, sentences are influenced by defendants’ pretrial status—whether they were detained or not, the prosecution’s position in the sentencing hearing, defendants’ representation, and their response to the charges. It is often difficult to control for all these factors and adequately show that it is the judge’s differentiated treatment that leads to what appears to be a harsher treatment of minorities.

This study attempts to overcome these difficulties in three ways. First, rather than asking whether there is an ethnic gap in judicial treatment, we ask whether there are systematic differences across the ethnic group of judges (see also Abrams et al. 2008). Such a comparison allows us to exploit the random assignment of cases to judges, which solves the selection problem that stems from unobservable case and suspect characteristics (Angrist & Pischke 2009; Shadish et al. 2002). This design enables us to assess whether ethnic disparities in judicial decisions are the result of unobservable case and suspect variables or due to judicial ethnic bias. The second way this study attempts to overcome the aforementioned methodological pitfalls is by directing attention to the earliest judicial decisions in the criminal process—the suspect’s first bail hearing—usually within 24 hours of his or her arrest by the police. At this stage, judges have to make decisions based on very limited information and those decisions are preceded, and thus potentially affected, by the least number of stages. Lastly, the Israeli bail hearing procedure requires the police to state the number of detention days they ask the court to order when applying for a detention order. This formally recorded number of requested days can serve to quantitatively control for the state’s position, which is one of the major factors in any decision in criminal cases. Thus, the setting of the first bail hearing in Israeli courts presents a unique opportunity for examining the effect of ethnicity on *judicial* decisions, beyond its potential effect on the state’s position in court.

Our findings show systematic evidence of ethnic in-group bias in the decision to detain or to release on bail. However, no ethnic in-group bias was found in the length of detentions. One inherent limitation of our study is that in the absence of a known ethnic-neutral baseline, we cannot formally determine whether our findings reflect discrimination against Arab suspects by Jewish judges, discrimination in favor of Arab suspects by Arab judges, or a mixture of both. Still, the study importantly and compellingly shows that judges are not blind to ethnicity. Notably, judges from different ethnic groups vary in their treatment of ethnic groups in a way that is consistent with ethnic in-group bias.

The article proceeds as follows. Section II reviews the literature on ethnic and racial bias in the criminal legal system and presents the general approach adopted in this study: (1) addressing ethnic in-group bias; (2) focusing on the earliest judicial decisions of the criminal process; and (3) utilizing the Israeli bail hearing process for qualitatively controlling for the state’s position. Section III describes the initial bail hearing process in Israel. Section IV presents the research design, data, and measurement. It also includes tests of randomization (of judge’s assignment to cases), provides descriptive statistics, and presents

the statistical method employed. Section V presents the results of both analyses—detention probability and detention length, and Section VI concludes with a discussion of the empirical and theoretical implications of the findings.

II. SCIENTIFIC BACKGROUND

In nearly every Western society, members of disadvantaged ethnic or racial minorities are overrepresented in jails and prisons (Tonry 1997). While the reasons for this observable fact are numerous, the possibility that judicial discriminatory practices are partly responsible is troubling. If otherwise equal cases are treated differently due to the ethnic identity of the participants, the legitimacy of the judiciary is undermined. After all, one of the major justifications for an independent judiciary is its ability to protect minorities from prejudice and stereotyping and assure equality before the law.

Thus, the effect of race and ethnicity on judicial decisions in criminal cases has attracted much attention in the literature. Numerous studies have examined whether disadvantaged minorities are treated differently, mostly in sentencing outcomes. Although many studies have found that minorities suffer from harsher sentences, the difficult question is not whether these disparities exist, but rather what their causes are. Are these disparities the result of discrimination in the criminal justice system? Do they result from differential criminal profiles of different ethnic groups? Do they stem from discrimination outside the criminal justice system? A central problem in providing a full answer to these questions draws on the lack of an available neutral baseline. Under these conditions, any statistical analysis of actual outcomes that employs a null hypothesis of equal outcomes across ethnic groups must assume control of all the relevant control variables. Hence, it is far from clear whether these disparate outcomes are the result of judges' ethnic and racial bias or whether they result from omitted case characteristics that are correlated with ethnic identity (Abrams et al. 2008). Kleck (1981) analyzed U.S. research on racial bias in sentencing and showed that the studies that appeared to find racial discrimination usually failed to adequately control for criminal record and other explanatory factors. More recent studies have led some commentators to conclude that there is little evidence of overt racial bias in sentencing (Klein et al. 1990; Sampson & Lauritsen 1997:362; Warner 2000). This does not mean that racial discrimination has disappeared; but, rather, if it exists, it is more subtle and harder to observe (Daly & Tonry 1997; Bushway & Piehl 2001).

Other analyses are less conclusive. For example, Chiricos and Crawford (1995) argue that there is sufficient evidence to infer direct racial discrimination in sentencing. Similarly, Spohn's (2000) survey of the literature led her to conclude that "earlier refutations of the discrimination thesis were premature." She reports that a majority of studies suggest that race impacts the incarceration decision, but less than one-quarter of them provide evidence that race affects sentence length. Still, these conclusions do not rule out the possibility that it is prosecutors, defense attorneys, or other nonjudicial actors who are responsible for the biased result.

Research findings in other Western countries are equally ambiguous. It seems that wherever the issue has been thoroughly investigated, “neither proponents of the claim that bias is a cause of disparities nor proponents of the claim that bias has no influence will find empirical evidence to prove their claims” (Tonry 1997). The research on sentencing of Jews and Arabs in the Israeli criminal justice system is no exception. Attempts to find judicial bias in sentencing have achieved mixed results, with several studies having found harsher treatment of Arabs (Cohen & Palmor 1985; Rattner & Fishman 1998; Fishman et al. 2006), while others have shown that evidence of differential treatment virtually disappears when pretrial detention and prosecutor position in the sentencing hearings were controlled for (Hassin & Kremnitzer 1993; Haj Yahia et al. 1994).

Why, despite the numerous studies that have been conducted, is it still so hard to find a clear view regarding the existence of racial and ethnic discrimination in judicial sentencing? The difficulty is that race and ethnicity are often correlated with other factors that have been found to significantly influence sentences, such as criminal record (Kleck 1981), pretrial detention (Demuth & Steffensmeier 2004b), unemployment (Chiricos & Bales 1991), court-appointed counseling (Holmes et al. 1996), crime type (Tonry 1995), aggravated circumstances (Kleck 1981), and so forth, and when these factors are controlled for, the independent effect of race loses all or most of its explanatory power in some of the studies. When so many racially correlated factors influence sentencing, it is statistically difficult to determine whether race in itself influences sentencing directly. Determining whether these factors receive unjustified weight because they mainly affect minorities is even harder.

Even when race and ethnicity has a direct effect on sentencing, it is still questionable whether discrimination should be attributed to the sentencing judges. As Bushway and Piehl (2001) explain, the sentence outcome is the result of the decisions of multiple actors interacting in one system. Sentence disparity might be the result of differentiated treatment by the police, prosecutors (Humphrey & Fogarty 1987), probation officers (Bridges & Steen 1998; Ben-Zion & Palmor 1985), defense attorneys, and other participants. In fact, it is likely that prosecutors have a more direct impact on racial disparities than do judges (Humphrey & Fogarty 1987; Davis 1998). Their sentencing recommendations have a substantial impact (Englich & Mussweiler 2001; Englich et al. 2005; Kremnitzer & Hassin 1993) but, even more importantly, prosecutors have direct control over the outcome through plea bargaining, which is the method of disposition in the vast majority of cases. If prosecutors are offering white defendants more lenient plea bargains than non-whites, the sentence disparity should be attributed to prosecutors, rather than to courts (Humphrey & Fogarty 1987). The existing research on sentencing has made virtually no attempt to distinguish between prosecutors’ and judges’ differentiated treatment of minorities and thus even if the studies that support the discrimination thesis are found more convincing than the contradicting studies, one cannot convincingly argue that this discrimination is the judges’ fault.

One way of isolating judicial ethnic bias is by examining decisions of judges from different ethnic groups. Although many studies have examined the effects of defendant race on judicial decision making, few studies have examined both judge and defendant race. Most of these studies did not find that defendant race had a different effect on the

decisions of black and white judges. Steffensmeier and Britt (2001) and Uhlman (1978) found that black judges sentence both black and white defendants somewhat more harshly than do white judges, but defendant race had similar effects on white and black judges. Similarly, two studies by Spohn (1990a, 1990b) found no significant effect of judge race on sentencing of white and black defendants. On the one hand, Welch et al. (1988) found that black judges treated black defendants more equally compared to white judges in the decision to incarcerate, and Johnson (2006) found that minority judges were less likely to incarcerate black and Hispanic offenders, but found no difference in the length of sentences. Price and Wolfers (2007), Parsons et al. (2007), and Larsen et al. (2008) found in-group racial bias in the decisions of officials in basketball and baseball leagues. Fishman et al. (2006) found that in Israeli courts, Jewish judges were less likely to impose a prison sentence on Jewish defendants (compared to Arab defendants), whereas the ethnic identity of the defendants did not seem to play a role in the decisions of Arab judges.¹ In contrast, Zussman and Shayo (2010) found evidence of ethnic in-group bias in Israeli small claims court rulings. Abrams et al.'s (2008) study of sentencing decisions in Cook County, Illinois found statistically significant between-judge variation in incarceration rates, although not in sentence lengths. All in all, the limited literature on the interaction between judge and defendant racial and ethnic background provides mixed evidence as to whether the harsher treatment of minorities in court results from judges' in-group bias.

This study differs from earlier works in three main respects, which together create a uniquely effective method of isolating judicial bias from other ethnicity-related factors. First, in line with Abrams et al. (2008), rather than asking whether there is an ethnic gap in judicial treatment, we ask whether there are systematic differences across the ethnic groups of judges. Such a comparison allows us to exploit the random assignment of cases to judges, which solves the selection problem that stems from unobservable case and suspect characteristics (Angrist & Pischke 2009; Shadish et al. 2002). This design enables us to assess whether ethnic disparities in judicial decisions are the result of unobservable case and suspect variables or due to judicial ethnic bias. If the former is the case, we should expect an overall difference in judicial treatment by suspects' ethnic group, but not a systematic variation in this difference across ethnic groups of judges, as random assignment ensures that each group receives the same case mix. However, if Arab and Jewish judges are found to systematically treat Arab and Jewish suspects differently, this would provide strong evidence for judicial in-group bias.

The second way this research attempts to overcome the aforementioned methodological pitfalls is by directing our attention to the earliest judicial decisions in the criminal process—the suspect's first bail hearing—usually within 24 hours of his or her arrest by the police. At this stage, judges have to make decisions based on very limited information and those decisions are preceded, and thus potentially affected, by the least number of stages.

¹The unique importance of Fishman et al.'s (2006) study is in the inclusion of the ethnic identity of the victim of the crime, and its interactions with the ethnic identity of perpetrators and judges, but this is less relevant to the current study. Note also that their interest in the effects of the victim's identity has led their study to look only at violent offenses (2006:74).

Each stage in the criminal process influences the result of the next stage. For example, it is well established that pretrial detention increases the probability of conviction and of imprisonment sentence (Rankin 1964; Clarke & Kurtz 1983; Frazier & Bishop 1985; Williams 2003). Similarly, the type of plea, the manner in which witnesses are questioned, the prosecution's positions at each step, and many other factors can be correlated with race and ethnicity and thus affect the decisions in the following stages. By examining the practice of judges at the very beginning of the process, rather than at the end, the effects of previous stages are minimized.

Focusing on early detention decisions also serves a theoretical consideration. It is often suggested that lack of information drives judges to rely on attributions linked to defendant race as a mechanism to reduce uncertainty (Albonetti 1991; Bridges & Steen 1998; Schlesinger 2005; Demuth & Steffensmeier 2004a). At the first bail hearing, which takes place only hours after the initial arrest, the judge has very limited information on the suspect and the offense. Thus, in this stage, ethnicity might have a different, more pronounced role (Demuth 2003). Only a few studies have examined the effect of race on pretrial detention decisions, without an attempt to isolate initial decisions and later ones. Like sentencing research, these studies do not clearly support or refute the discrimination hypothesis. Several studies found indication of discrimination (Ayres & Waldfogel 1994; Bynum 1982; Patterson & Lynch 1991; Petee 1994). Yet even in this relatively initial stage, some ethnically or racially correlated factors can influence the result and thus the result is heavily influenced by the variables that were unavailable to the researchers.² The approach adopted in this study thus reduces these potential unobservable case variables by focusing on pretrial detention decisions, and overcomes the remaining problem by looking at the variation across the ethnic groups of randomly assigned judges.

Lastly, the Israeli bail hearing procedure allows us to control for the overall state's position. When applying for a detention order, the police are required to state the number of detention days they request the court to order. This written number of requested days serves in our analyses to quantitatively control for the state's position, which is one of the major factors in any decision in criminal cases. To the best of our knowledge, none of the existing pretrial detention studies have used such a control. Disregarding the state's position in bail hearings is as problematic as disregarding plea bargaining in sentencing. To the extent that the state's position is systematically associated with ethnic characteristics, any potential disparity in the result might be attributed to the state and not to the judge. Obviously, the state's agents might also be biased. Yet controlling for the state's position in the analyses enables us to infer *judicial* bias from racially differential treatment in the outcome of the judicial decision, separately from other sources of bias in the criminal justice system. Thus, the setting of the first bail hearing in Israeli courts presents a unique opportunity for examining the effect of ethnicity on judicial behavior, isolating judicial bias from other relevant and irrelevant factors already considered by the state.

²Ayres and Waldfogel (1994) concede that although their analysis controls for many variables, they might still be missing one or more omitted variables that may account for the disparate outcome.

III. THE INITIAL BAIL HEARING IN ISRAEL

Israeli criminal procedure provides an excellent laboratory for examining the effect of ethnicity on pretrial decisions. First, it operates within a divided society in which criminal suspects and judges include both Arabs and Jews. Second, the pretrial detention hearing is the earliest stage in the criminal process in which a judicial decision is involved. Third, the prosecution's position is quantitatively presented in the hearing by the requested length of detention and, finally, judges and public defenders are randomly allocated for the different cases.

In this system, Arabs play much the same role as disadvantaged minorities elsewhere. While there are internal ethnic or faith-based divisions within the Jewish and the Arab groups, members of each group usually see themselves as part of one single ethnic and national identity. Since the establishment of Israel in 1948, a fragile and contested coexistence between the Jewish majority and the Arab minority has gradually been developed (for a recent review of this relationship, see Samoooha 2005). Hostile undercurrents in these intergroup relations are further augmented by the ongoing Israeli-Arab conflict. It is therefore not surprising that these social realities permeate the courts. Indeed, a sizable number of studies provide consistent evidence that there is a wide disparity in sentencing of Jews and Arabs in the Israeli justice system (Cohen 1985; Cohen & Palmor 1985; Hassin & Kremnitzer 1988; Kretzmer 1990; Rattner & Fishman 1998; Fishman et al. 2006). Like African Americans in the United States or minority groups in the United Kingdom, Arabs are overrepresented in prisons; they suffer from negative stereotypes and are more likely to be socially associated with severe crimes (Fishman et al. 1987). However, as in other countries, a strand of the literature suggests that when controlling for other major factors such as criminal record, pretrial detention, and prosecutor and sentencing requests, little evidence of judicial discrimination remains (Hassin & Kremnitzer 1993; Haj Yahia et al. 1994).

As in many other jurisdictions, a suspect in Israel has to be brought before a judge within 24 hours of the arrest, unless a police officer has decided to release him or her earlier. The court can then detain the suspect for up to 15 days, though it usually orders a much shorter detention period. In the case that the suspect is detained, the police can ask the court to order the continuation for up to 15 days from time to time, as long as the total detention period does not exceed 30 days prior to the issuance of an indictment. After an indictment is issued, the case is handled by the prosecution, which can either agree to release the defendant on bail or request the court to remand the defendant for the full duration of the trial.

In the pretrial hearing, the state is represented by a police detective (a prosecutor represents the state only after an indictment is issued). In the application, the police officer asks the court either to release the suspect on bail or to detain him or her. In the latter case, the officer also states how many days of detention are requested. The court reads the request and hears the suspect or his or her attorney, and then decides whether to detain the suspect and, if so, for how long. The factors considered at this stage are the suspect's dangerousness, the risk of flight, the need to prevent disruption of proceedings, and investigation needs.

The suspect is brought before a magistrate court judge, who has been designated in advance to sit on all the bail hearings on that specific day. The public defender's office sends a lawyer to represent all suspects who do not hire private lawyers.³ Hence, judges and public defenders are randomly allocated for the different cases. With this in mind, we turn to the method of our study.

IV. RESEARCH DESIGN

Data were drawn from the entire list of pretrial bail hearings taken on Fridays and Saturdays during 2004 at the Haifa, Nazareth, and Acre magistrate courts ($N = 1,852$).⁴ Fridays and Saturdays were selected for two reasons. First, while during working days particular judges handle this type of hearing, all the judges serve on prearranged rotation duty in these hearings over the weekend. This provides a "naturally occurring" process of random allocation of judges to cases. Second, during these days, nearly all hearings deal with first detention decisions, that is, regarding suspects that have first been apprehended during the previous 24 hours, while a larger proportion of continuation of detention decisions typifies other days. The number of judges in the sample is 59, out of whom 15 are Arabs and 44 are Jews. It should be noted that the three courts chosen include the majority (78.9 percent) of the Arab magistrate judges in Israel.⁵ Table 1 provides initial evidence consistent with a random assignment of judges to cases, showing that none of the case variables are related to the ethnic identity of the judge. Appendix A also shows that none of these variables have any power in explaining the assignment of judges (Arabs or Jews) to cases.

The analysis included two stages. In the first stage, the judicial decision was analyzed as a dichotomous variable—release or detain. The second stage analyzed the decision as a count variable based on the length of detention in days, from 1 to 15. Only those cases in which agreement was not reached among the parties regarding release or the length of detention, thereby resolved solely by a judicial decision, were included in the analyses.⁶ Additionally, cases in which the police request was for bail (zero detention days) were

³At this stage, suspects have a right to public representation only if they are found unable to hire a lawyer of their own, but since the hearing takes place soon after the arrest, it is impossible to inquire about a suspect's financial situation and thus, in practice, every suspect is offered a public defender.

⁴The weekend in Israel includes Friday and Saturday, while Sunday is a workday.

⁵A total of 19 Arab magistrate judges (excluding labor and family courts, which do not handle detention decisions) (based on the official Judicial Authority website: <<http://www.court.gov.il>>).

⁶Agreed decisions are the result of three possible processes. One-third of the agreed decisions are determined before the hearing, and the judge simply approves the parties' agreement. The rest of the agreements are formed when the judge sends the parties to deliberate (typically in the corridor) and later approves their agreement, or suggests a possible outcome of the agreement and suggests to the parties to reach a proximate agreement. The exact proportions of these two situations are unknown. Clearly, consensual decisions follow a different dynamic from judicial decisions, and they are studied more closely by the authors elsewhere.

Table 1: Randomization Tests

	<i>Arab Judge</i>	<i>Jewish Judge</i>	<i>p Value</i>
Arab suspect	45.9%	44.0%	0.579
Police request (in days)	5.007	5.005	0.987
Low severity	26.0%	24.2%	0.551
Mid-low severity	44.8%	46.3%	0.658
Mid-high severity	10.8%	10.7%	0.952
High severity	18.4%	18.8%	0.889
Violent offenses	24.5%	25.3%	0.796
Illegal residence	2.5%	4.4%	0.158
Sexual offenses	3.6%	3.8%	0.895
Property offenses	23.8%	26.8%	0.324
Domestic violence	22.0%	19.6%	0.386
Drug-related offenses	13.0%	11.7%	0.547
Traffic violations	2.5%	2.2%	0.753
Public defender	60.7%	65.3%	0.158
Unrepresented	9.4%	9.3%	0.981
Prosecutor's statement	9.8%	7.0%	0.134
Day (Saturday)	36.5%	37.5%	0.762

N = 1,230.

omitted from the data set, as in these cases the release was not in dispute.⁷ Thus the effective sample used for the analyses included 1,230 decisions.

We coded judges as Arabs or Jews based on their names and biographic information found on the Israeli Court System's official website. The same sources were used to code judge gender.⁸ Suspects were coded as Arabs or Jews based on their names as they appear in the protocol of the hearing. Other variables coded were the police request (1–15 days), whether the hearing took place on Friday or Saturday, and severity of the most severe offense mentioned in the detention procedure. Coding of offense severity followed a four-scale ordinal variable that represented the maximum statutory penalty for the offense—(1) less than three years (misdemeanors under Israeli law); (2) three to seven years; (3) seven to fifteen years; (4) over fifteen years. Based on earlier research (Hassin & Kremnitzer 1993), a dummy variable indicating the identity of the court—Nazareth and Acre—was included (neither indication represents cases from the Haifa court). The type of representation provided in each case was coded (private, public defender, or unrepresented), and the type of offense was coded following eight categories. In some of the cases, the police investigation is concluded by the time of the bail hearing. In these cases, detention is possible only if the prosecutor declares that the state intends to issue an

⁷Indeed in *all* these cases suspects were not detained by judges.

⁸See <http://www.court.gov.il>. The website also includes pictures of the individual judges. We can safely say that the ethnic identification of the judges is highly reliable. In the few cases in which we had any doubt, we verified the identity by contacting people who knew the judge in question. In one particular case we found that a certain judge was the son of an Arab father and a Jewish mother (which is quite rare in Israeli society). His decisions (only nine) were excluded from the data set.

Table 2: Summary Statistics

	<i>Arab Suspect</i>		<i>Jewish Suspect</i>		<i>Difference</i>
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>	
Police request (in days)	5.106	1.806	4.925	1.678	0.181*
Low severity	0.260	0.439	0.235	0.425	0.025
Mid-low severity	0.429	0.495	0.484	0.500	−0.055*
Mid-high severity	0.136	0.342	0.085	0.279	0.051***
High severity	0.176	0.381	0.196	0.397	−0.020
Violent offenses	0.273	0.446	0.234	0.424	0.039
Illegal residence	0.082	0.275	0.006	0.076	0.077***
Sexual offenses	0.037	0.188	0.038	0.191	−0.001
Property offenses	0.258	0.438	0.264	0.441	−0.005
Domestic violence	0.167	0.373	0.230	0.421	−0.063***
Drug-related offenses	0.095	0.294	0.139	0.346	−0.044**
Traffic violations	0.027	0.164	0.019	0.137	0.008
Public defender	0.546	0.498	0.719	0.450	−0.173***
Unrepresented	0.097	0.296	0.091	0.287	0.006
Prosecutor's statement	0.079	0.270	0.075	0.263	0.004
Day (Saturday)	0.366	0.482	0.377	0.485	−0.012
Nazareth	0.291	0.455	0.268	0.443	0.024
Acre	0.403	0.491	0.212	0.409	0.191***
Release	0.234	0.424	0.300	0.458	−0.065**
Detention length [†]	2.890	1.322	2.816	1.300	0.074
Judge ethnicity (Arab)	0.233	0.423	0.219	0.414	0.013
Judge gender (male)	0.676	0.468	0.639	0.481	0.037

N = 1,230; †*N* = 897; *, **, and *** denote differences that are statistically significant at 10 percent, 5 percent, and 1 percent, respectively.

indictment and request that the defendant be detained for the full duration of the proceedings. Since the official legal justification for detention in these cases is slightly different, a variable for such prosecutorial declaration was also included in the analysis. Table 2 provides a list of the variables used in our analysis, as well as a comparison of the mean values between Arab and Jewish suspects. We do not have data on the defendants' prior record as this information was not included in the files. However, the control for the state's position in the hearings is likely to incorporate most of the relevant factors for the bail hearing, including prior record. More importantly, the random assignment of cases to judges eliminates the risk that prior criminal record will be able to account for the behavioral differences between Jewish and Arab judges.⁹

Table 2 shows that the police ask for slightly more detention days for Arab suspects, compared to Jewish suspects (5.106 vs. 4.925 days); Arab suspects do not differ from Jewish suspects in the extreme ranges of offense severity, yet they appear to be less likely to be

⁹The omission of this variable (like any other) may act to reduce the explained variance in the analyses, but in the context of an experimental design this does not derogate from reliability of the estimate of the randomly assigned causal effect—in our case, the ethnic identity of the judge (see King 1991:1050–51).

detained by the police for mid-low-severity offenses compared to Jewish suspects (0.429 vs. 0.484), and more so for mid-high-severity offenses by a similar margin (0.136 vs. 0.085). Unsurprisingly, Arab suspects are much more involved in immigration offenses (0.082 vs. 0.006);¹⁰ but are less frequently arrested for domestic violence (0.167 vs. 0.230) and drug-related offenses (0.095 vs. 0.139) compared to Jewish suspects. Arab suspects are less frequently represented by the public defender's office (0.546 vs. 0.719), and significantly exceed the proportion of Jewish suspects in the magistrate court of Acre (0.403 vs. 0.212). As to the dependent variables, Arab suspects have a lower proportion of releases compared to Jewish suspects (0.234 vs. 0.300), yet they do not appear to differ in their mean length of detention (2.890 days vs. 2.816 days).

A. Statistical Analysis

To account for the data structure, the following analyses use hierarchical modeling procedures, designed to account for the nested nature of detention decision within judges.¹¹ The models examining release versus detention were estimated with logistic and multilevel logistic regressions.

Detention length provides a number of complications that should be accounted for in the analysis. One may consider detention length as including releases as "zero-value detentions," or exclude releases from the range of detention length as categorically distinct. We tend to subscribe to the later conceptualization. Moreover, even if one considers release as a zero-value case of detention, this value occupies the lower limit of the range, potentially misrepresenting cases with theoretically varying values of underlying judicial valuation. For example, such cases may include those cases that only just fall short of sanctioning detention; those that clearly should be released; and model citizens that should not have been arrested by the police in the first place. Thus including zero detention cases as another value of the detention variable would violate the validity of lower values of this measure and potentially inflate the number of zeros in such a variable. More formally, a strictly positive (excluding "releases") variable of detention length is expected to follow a Poisson distribution,¹² and one that includes zero detention length is expected to follow a zero-inflated Poisson distribution.¹³ Figure 1 graphically presents the histograms of detention length against the corresponding Poisson distributions. The two graphs confirm the theoretical expectations. It appears that the strictly positive variable approximates a Poisson distribution, while the variable that includes "releases" as zero-value detentions exhibits an

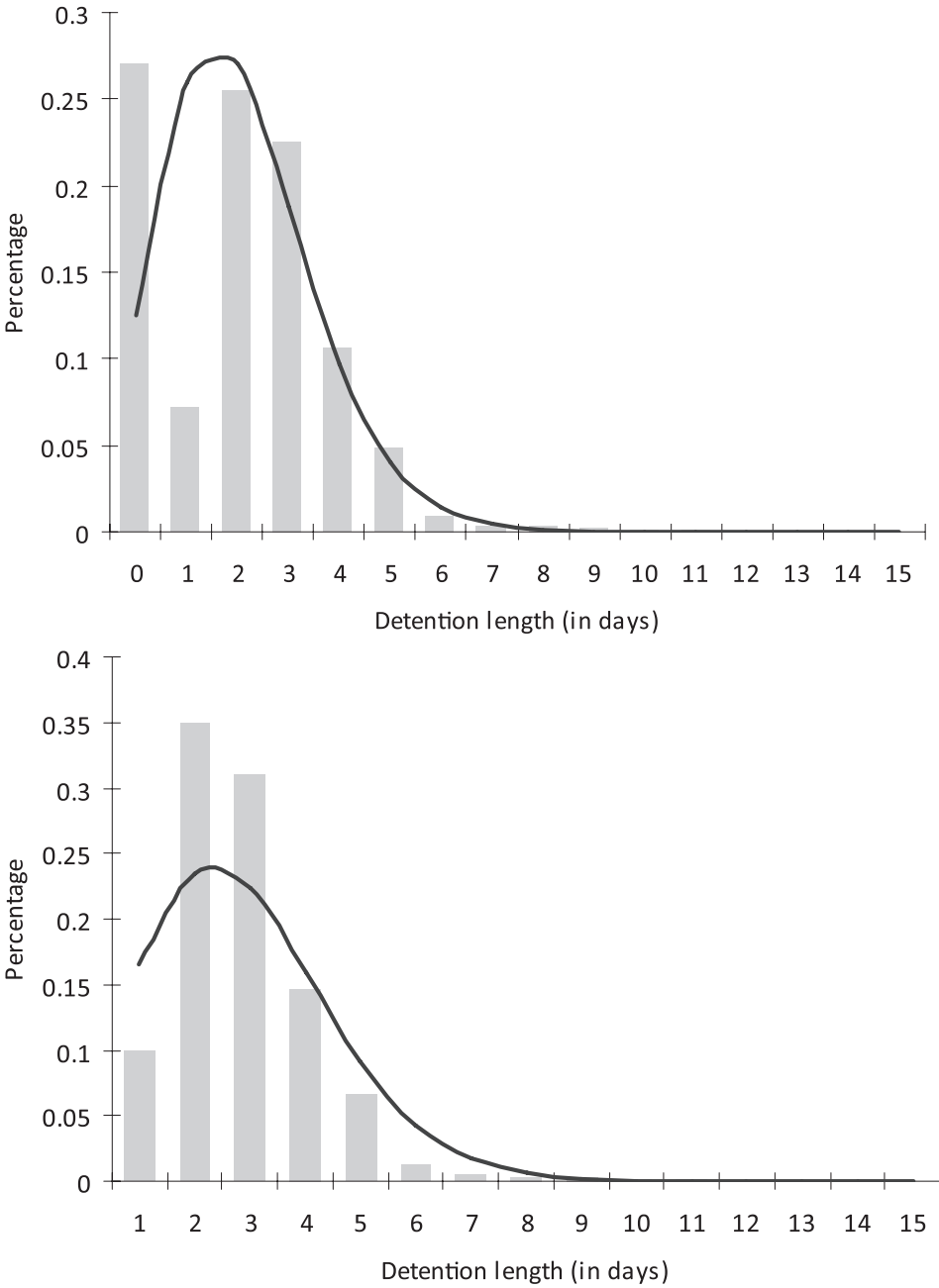
¹⁰This offense is almost entirely committed by Arab Palestinians from the West Bank and Gaza.

¹¹See also Johnson (2006). For a concise review of the method, see Steenbergen and Jones (2002).

¹²The mean detention length of this variable is 2.85 days and the variance is 1.716; the overdispersion parameter alpha is insignificantly different from zero ($p = 1.00$).

¹³The mean detention length of this variable is 2.08 days and the variance is 2.86; the overdispersion parameter alpha is insignificantly different from zero ($p = 1.00$).

Figure 1: Actual distribution of detention length including releases ($\lambda = 2.08$ days, $N = 1,230$) (upper graph); and excluding releases ($\lambda = 2.85$ days, $N = 897$) (low graph); and corresponding Poisson distributions.



excessive number of zeros.¹⁴ For these theoretical and empirical reasons, our main analyses have used the strictly positive variable of detention length, and the models examining detention length were estimated with Poisson and multilevel Poisson regressions.

Notwithstanding this, three sets of robustness tests follow the main analyses. The first is intended to assess the robustness of the findings regarding detention length under alternative distributional assumptions. It includes analyses of detention length with linear and multilevel linear regressions. The second and third sets of analyses gauge the possibility of selection bias in estimating detention duration as a result of including only those cases in which suspects were detained. These employ the zero-inflated Poisson regression (ZIP) and the Heckman two-step selection model. The detailed analyses are reported in the following section.

V. RESULTS

A. *The Effect of In-Group on the Probability of Release*

Table 3 presents the results of four logistic regression analyses with the log-odds of release as dependent variable. The second column of each model reports the odds-ratios after exponential transformations of the regression coefficients. We begin with a bare-bones specification in Model 1, which is a basic difference-in-difference specification. Models 2 and 3 add controls for observable judge, court, and case variables.¹⁵ Model 4 presents a multilevel analysis thereby controlling for unobserved variation across individual judges.

The primary focus of our analysis is the interaction term *Arab judge* \times *Arab suspect_{ij}*, which indicates whether the effect of the ethnic identity of the suspect on the probability of release is different between Arab and Jewish judges. Across all these specifications, we find that this difference is significant and positive. Given that the assignment of judges to cases is random, this statistical association enables us to infer that ethnic in-group-ness increases the likelihood of release. To demonstrate this finding, Figure 2 presents the predicted probabilities of release for Arab and Jewish suspects across Arab and Jewish judges, holding all other variables at their modal values.¹⁶ The significant difference between the effect of the suspect's ethnic identity on Arab and Jewish judges, inferred from the significant interaction term, is graphically depicted by the higher probability of release of Arab suspects for Arab judges, compared to the higher probability of release of Jewish suspects

¹⁴Indeed, goodness-of-fit χ^2 for the full Poisson regression model with strictly positive variable suggested a good fit (292.18, $p = 1.00$), while this was not the case for the variable that includes releases (1649.33, $p < 0.001$).

¹⁵Initial VIF tests of multicollinearity have shown a mean VIF of 2.20, yet three variables exceeded a VIF value of 4.0, with "property offenses" having the highest value (5.06). Excluding "property offenses" from the analysis has led to a mean VIF of 1.59 and a maximum level of 2.77 (high severity). To avoid multicollinearity and given that "property offenses" were not found to have a significant association with any of the dependent variables, nor with the ethnic identity of suspects and judges, they were excluded from the final analyses.

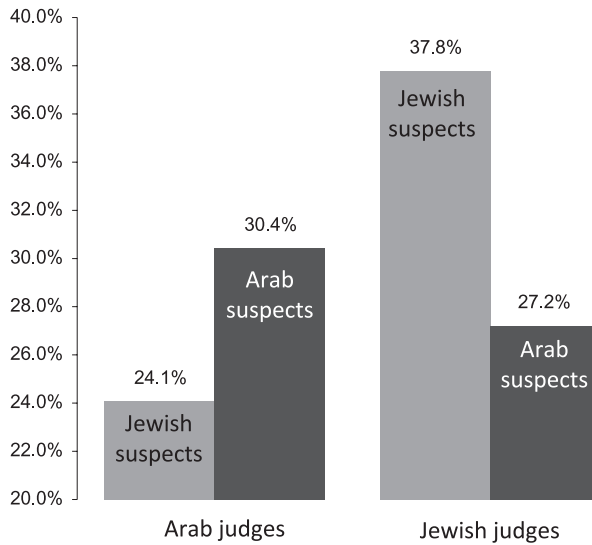
¹⁶Male judge, police request = 5 days, Haifa Magistrate Court, mid-low severity of offense, violent offense, representation by a public defender, on a Friday.

Table 3: Estimating the Effects of Ethnic In-Group on Release Rates—Logistic and Multilevel Logistic Regressions

	1		2		3		4	
	b	O.R.	b	O.R.	b	O.R.	b	O.R.
Arab judge × Arab suspect	0.631* (0.325)	1.880	0.810** (0.346)	2.248	0.832** (0.351)	2.298	0.806** (0.360)	2.239
Arab judge	−0.618*** (0.222)	0.539	−0.757*** (0.241)	0.469	−0.796*** (0.244)	0.451	−0.651** (0.298)	0.522
Arab suspect	−0.459*** (0.148)	0.612	−0.430*** (0.160)	0.651	−0.476*** (0.169)	0.621	−0.485*** (0.173)	0.616
Judge gender (male)			0.532*** (0.154)	1.702	0.520*** (0.158)	1.682	0.428** (0.206)	1.534
Police’s request			−0.415*** (0.060)	0.661	−0.434*** (0.062)	0.648	−0.446*** (0.064)	0.640
Court (Nazareth)			−0.964*** (0.191)	0.381	−0.946*** (0.197)	0.388	−0.952*** (0.243)	0.386
Court (Acre)			−0.256 (0.164)	0.774	−0.280 (0.170)	0.756	−0.342 (0.244)	0.710
Mid-low severity			−0.142 (0.161)	0.868	−0.310 (0.213)	0.733	−0.298 (0.216)	0.742
Mid-high severity			−0.628** (0.282)	0.534	−0.726** (0.298)	0.483	−0.733** (0.303)	0.480
High severity			−0.708*** (0.234)	0.492	−0.835*** (0.303)	0.434	−0.813*** (0.308)	0.444
Violent offense					−0.104 (0.212)	1.109	0.117 (0.216)	1.124
Illegal residence					−0.008 (0.390)	0.992	−0.034 (0.397)	0.967
Sexual offense					0.487 (0.433)	1.627	0.524 (0.439)	1.689
Domestic violence					0.584*** (0.197)	1.793	0.581*** (0.200)	1.788
Drug-related offense					0.203 (0.346)	1.225	0.165 (0.351)	1.179
Traffic violation					1.108** (0.463)	3.027	1.162** (0.468)	3.196
Public defender					−0.378** (0.169)	0.685	−0.373** (0.173)	0.689
Unrepresented					−0.374 (0.266)	0.688	−0.364 (0.272)	0.695
Prosecutor’s declaration					−0.787** (0.318)	0.455	−0.862*** (−.325)	0.422
Day (Saturday)					−0.106 (0.150)	0.899	−0.119 (0.156)	0.888
Constant	−0.727*** (0.092)		1.405*** (0.322)		1.789*** (0.409)		1.858*** (0.431)	
Ln(sigma u)							−1.885 0.679	
Likelihood-ratio test of ρ							4.72**	
N	1,229		1,229		1,229		1,229	

*, **, and *** denote significance levels of 10 percent, 5 percent, and 1 percent, respectively.

Figure 2: Predicted probabilities of release across the ethnic identity of judges and suspects (at the modal case).



for Jewish judges. At the modal situation, Arab suspects are 6.3 percent more likely than Jewish suspects to be released by an Arab judge, while Jewish suspects are 10.4 percent more likely than Arab suspects to be released by a Jewish judge.

It should be noted that the significant *Arab judge* \times *Arab suspect*_{ij} interaction term refers to the difference between the treatment of Arab and Jewish suspects in the comparison between Arab and Jewish judges (as the null hypothesis is no difference in the treatment across judge groups). However, analyses of release rates *within* each group of judges involve comparisons between the actual disparate rates of release (e.g., 24.1 percent and 30.4 percent) and an *equal* release rate of Arab and Jewish suspects. Despite our attempt to control for many case, judge, and court variables, the true ethnically-neutral baseline of release probability is not known and not necessarily equal across suspects' ethnic groups. Thus within-judges'-group comparisons cannot be used to infer whether Jewish, Arab, or both groups of judges discriminate. Yet regardless of the ethnically-neutral baseline of release probability, the differences in the treatment of Arab and Jewish suspects by the two groups of judges clearly show that ethnic identity plays a role in these decisions, in a pattern that is consistent with in-group bias.

As could be expected, the police request was found to be negatively associated with the probability of release.¹⁷ The severity of the offense is also negatively associated with

¹⁷The predicted probability of release for the case in which the police request four detention days and other variables are at their modal value is 48.7 percent. An increase of one day in the police's request reduces the probability of release to 37.8 percent.

release probability (45.0 percent, 37.8 percent, 28.2 percent, and 26.6 percent for low, mid-low, mid-high, and high severity levels, respectively).¹⁸ Interestingly, male judges are more likely to release suspects on bail (37.8 percent vs. 28.4 percent). Suspects in the Nazareth Magistrate Court are less likely to be released (19.0 percent vs. 30.2 percent at Acre, and 37.8 percent at Haifa), and so are suspects who are represented by a public defender (37.8 percent) or are unrepresented (38.0 percent) compared to those represented by a private attorney (46.9 percent). Within the types of offenses, domestic violence and traffic violations were found to have a significant (positive) association with the probability of release (49.2 percent and 63.3 percent, respectively, compared to the modal (violent) offense—37.8 percent). It appears that accounting for the multilevel structure of the data resulted in some changes in the regression estimates. This result is consistent with the significant likelihood-ratio test of rho ($p = 0.015$) and suggests that interjudge variance in release decisions is substantive and statistically significant.

B. The Effect of In-Group on the Length of Detention

Table 4 presents the results of four Poisson regression analyses with detention length as the dependent variable. This analysis was based on the subset of decisions to detain (no releases) ($N = 897$). In this analysis, we also begin with a bare-bones specification in Model 5. Models 6 and 7 add controls for observable judge, court, and case variables, and Model 8 presents a two-level analysis, thereby controlling for unobserved variation across individual judges.

No significant *Arab judge* \times *Arab suspect_{ij}* interaction was found for any of the models in Table 4 nor any main effect of judge or suspect ethnicity on the length of detention. The police request was found to be significantly and positively associated with the length of the detention ordered, that is, every additional day of detention requested by the police resulted in an increase of 12.7 percent in detention length (e.g., an increase from 5 to 6 days requested by the police is associated with an increase of 0.635 days of detention). Although male judges were found to have a greater propensity to release suspects (compared to female judges), when they did decide to detain, they tended to order 10.3 percent longer detentions. Drug-related offenses were found to result in 20.6 percent longer detentions, and detentions on Saturdays were 13.6 percent shorter than those ordered on Fridays. It appears that accounting for the multilevel structure of the data resulted in negligible differences in the regression estimates. This result is consistent with the insignificant likelihood-ratio test of alpha ($p = 0.499$) and suggests that interjudge variance in determining detention length is statistically insignificant and practically negligible.

C. Robustness Tests

To assess the robustness of the findings regarding detention length, these relationships were also estimated with linear and multilevel linear regressions, which are reported in Appendix B. Again, we begin with a bare-bones specification in Model 9, adding controls for observable judge, court, and case variables in Models 10 and 11, and presenting a

¹⁸This and the following comparisons were computed when all other variables were held at their modal values.

Table 4: Estimating the Effects of Ethnic In-Group on Detention Length—Poisson and Multilevel Poisson Regressions

	5	6	7	8
	b	b	b	b
Arab judge × Arab suspect	0.077 (0.092)	0.080 (0.093)	0.083 (0.093)	0.083 (0.093)
Arab judge	−0.029 (0.063)	−0.051 (0.065)	−0.053 (0.065)	−0.053 (0.065)
Arab suspect	0.007 (0.045)	−0.021 (0.047)	−0.015 (0.049)	−0.015 (0.049)
Judge gender (male)		0.118*** (0.043)	0.099** (0.044)	0.099** (0.044)
Police’s request		0.119*** (0.009)	0.120*** (0.010)	0.120*** (0.010)
Court (Nazareth)		−0.017 (0.048)	−0.028 (0.050)	−0.028 (0.050)
Court (Acre)		−0.037 (0.052)	−0.067 (0.053)	−0.067 (0.053)
Mid-low severity		−0.038 (0.055)	−0.003 (0.067)	−0.003 (0.067)
Mid-high severity		0.060 (0.071)	0.042 (0.076)	0.042 (0.076)
High severity		0.024 (0.064)	−0.071 (0.079)	−0.071 (0.079)
Violent offense			0.062 (0.059)	0.062 (0.059)
Illegal residence			0.028 (0.128)	0.028 (0.128)
Sexual offense			0.113 (0.108)	0.113 (0.108)
Domestic violence			0.011 (0.064)	0.011 (0.064)
Drug-related offense			0.187** (0.087)	0.187** (0.087)
Traffic violation			−0.015 (0.186)	−0.015 (0.186)
Public defender			−0.026 (0.048)	−0.026 (0.048)
Unrepresented			0.006 (0.077)	0.006 (0.077)
Prosecutor’s declaration			−0.009 (0.078)	−0.009 (0.078)
Day (Saturday)			−0.146*** (0.044)	−0.146*** (0.044)
Inverse Mills ratio				
Constant	1.042*** (0.031)	0.359*** (0.074)	0.395*** (0.098)	0.395*** (0.098)
Ln(alpha)				−18.864 (859.545)
Likelihood-ratio test of α				6.9·10 ^{−6}
N	896	896	896	896

*, **, and *** denote significance levels of 10 percent, 5 percent, and 1 percent, respectively.

two-level linear analysis in Model 12. The findings were generally consistent with those found in the Poisson models.¹⁹ More specifically, no significant *Arab judge* \times *Arab suspect_{ij}* interaction was found for any of the models nor any main effect of judge or suspect ethnicity on the length of detention.

The second and third sets of analyses are intended to assess the possibility of selection bias in estimating detention duration as a result of including only those cases in which suspects were detained. First, the zero-inflated Poisson regression (ZIP) was estimated using the entire set of decisions. The count model for detention length included the dependent variables regarding the ethnicity of the judges and suspects, judge gender, the police's request, offense type, and hearing day (Friday/Saturday), omitting variables that were not found to be significantly associated with detention length in Model 8. The model predicting release included the dependent variables concerning the court's location, offense severity, and representation category. The analysis is reported in the left column in Appendix C. The significant Young test result ($p < 0.001$) indicates that the ZIP model fits these data better than the standard Poisson. The estimates in the count model are consistent with those found in Models 7 and 8, and the estimates predicting release conform to those found in Model 4 (with the exception of the prosecutor's declaration).

Next, the Heckman two-step selection model (Heckman 1979) was estimated using the entire set of decisions. The allocation of dependent variables to the two models is identical to the one used in the ZIP analysis. This separation between the dependent variables follows Puhani (2000) in order to decrease multicollinearity in the data.²⁰ The results of this model are reported in the center column in Appendix C. Finally, the inverse Mills ratio calculated by the Heckman two-step model is included as an additional dependent variable in the multilevel Poisson regression (see Green 1994). Both linear and multilevel Poisson models with Heckman two-step selection correction conform to the results of Models 7 and 8. We thus conclude that our findings in the analysis of the effects of ethnic in-group on detention length appear to be robust.

VI. DISCUSSION

This is the first study to utilize a natural randomized experiment to assess the effect of suspect ethnic identity on pretrial detention decisions across the ethnic identity of the deciding judges. We have found support for in-group bias in the likelihood of release, while we have found no support for such bias in the decision on the length of detention.

Comparing these findings to those of previous studies in the context of sentencing in the United States presents a mixed result. They are contrary to the findings of Uhlman

¹⁹One exception is the finding of a significant association between "violent offense" and a positive increase in detention length, which was found in the multilevel linear regression analysis (Model 12) and was not found in the Poisson analyses.

²⁰The full model R^2 of the regression of the inverse Mills ratio on the regressors of detention length is 0.67; mean VIF = 1.61, maximum VIF = 2.80.

(1978), Spohn (1990b), and Steffensmeier and Britt (2001), but conform to the findings of Welch et al. (1988) and, particularly, Johnson (2006), who found that minority judges are less likely to incarcerate black and Hispanic offenders, yet found no difference in the length of sentences. Our findings also coincide with Abrams et al. (2008), who found statistically significant between-judge variation (yet not specifically across judge ethnic identity) in incarceration rates, although not in sentence lengths. In the Israeli sentencing context, our findings agree with those of Fishman et al. (2006), who found that Jewish judges were less likely to impose a prison sentence on Jewish defendants (compared to Arab defendants), whereas the ethnic identity of the defendants did not seem to play a role in the decisions of Arab judges. This inconsistent level of agreement between our findings and previous studies is not surprising given the conflicting results in the literature. We would, however, note the fact that both U.S. and Israeli studies are included in those studies that are similar in their findings to ours, suggesting that these results are not limited to the Israeli context.

A. Judicial Ethnic In-Group Bias as Implicit Bias

How should we account for the evidence for ethnic in-group bias in the decision to release or detain and the lack of such bias in the decision on the length of detention, which also mirrors similar findings in the context of sentencing (Johnson 2006; Abrams et al. 2008)? The main difference between these two decisions is in the characteristic of the response alternatives—release or detain—as opposed to a range of detention lengths of from 1 to 15 days. Judges are required to consider a series of varied details pertaining to the criminal behavior, its particular circumstances and consequences, future potential risks that the suspect poses if released, and a range of legal rules that apply to all these facts. While the first decision (detention/release) requires them to transform their overall impression from this elaborate set of information to one of two possible responses, the second one (detention duration) offers a more varied range of response choice.

Under these conditions, one possible interpretation of the findings is that judicial ethnic in-group bias is an implicit bias. An implicit attitude is defined as an unconscious association between a target (e.g., an Arab suspect) and a given attribute (Bertrand et al. 2005; Greenwald & Krieger 2006). Such attitudes exist regardless of whether they are aligned with a person's explicit (conscious) attitudes, and thus affect judgment and behavior in unintentional and often unconscious ways. Rachlinski et al. (2009) have found that judges harbor similar implicit racial biases to those found in the general population, yet they can suppress their implicit racial bias from influencing their decisions in some circumstances. Although both implicit and explicit attitudes are at work in determining judgment, choice, and behavior, it appears that implicit bias may have greater influence in less-controllable behavioral outcomes. Notably, three conditions are conducive to the influence of implicit bias: inattentiveness to task, time pressure or other cognitive load, and task ambiguity (Bertrand et al. 2005).²¹ Although the first two conditions are

²¹One possible explanation for this effect of ambiguity on the expression of implicit bias is the process of "attribute substitution," in which people confronted with a difficult question often answer an easier one instead, usually without being aware of the substitution (Kahneman & Frederick 2002; Jolls & Sunstein 2006).

equal across the two judicial choices under question here, the third appears to be relevant.

Ambiguous tasks are those that do not provide a simple formula for their completion.²² Deciding between releasing and detaining a suspect based on the information presented before a judge in a bail hearing forces him or her to make a choice between two very distinct outcomes even when the relevant considerations suggest an equivocal situation. Such a task thus imposes (at least in some cases) an implicit “rounding up” or “down” of case information evaluation in order to make a choice. This implicit “rounding” of evaluation is unguided and thus ambiguous. On the other hand, the decision on the length of a detention provides a range of ranked responses from 1 to 15 days, which facilitates matching a measured response to the case’s specific circumstance. We therefore suggest that the finding of ethnic in-group bias in the decision to release or detain but not in the length of detentions can be understood as resulting from implicit rather than explicit bias. This bias is variably manifested in the judicial decisions based on their level of task ambiguity.²³ Further research is needed in order to assess this potential explanation, both in the context of detentions and in sentencing.

B. From In-Group Bias to Ethnic Discrimination

The difficulties in studying ethnic discrimination in the legal system, discussed in the introduction to this article, are mainly due to potential unobserved variables that entail uncertainty as to the true neutral baseline for treating members of different ethnic groups. Looking at *absolute* differences in the treatment of ethnic group members requires us to control all relevant variables in order to be able to equate the null hypothesis of “zero difference” with the true neutral treatment. Indeed, differences in controls in previous studies, in Israel as in other countries, have resulted in mixed findings. Alternatively, replacing “ethnic bias” with “ethnic in-group bias” as the concept under investigation diverts our attention to *relative* difference in the (apparent) differential treatment across the ethnic identity of judges, taking advantage of the random allocation of judges to cases, and provides an opportunity to avoid this problem. In this study, the null hypothesis is *not* that the rate of release or length of detention of members in the two ethnic groups is equal, but that the treatment of these two groups (equal or not) is the same across the ethnic group of judges. Such a null hypothesis conforms to the normative expectations from judges in a system that upholds equality before the law, and its manifestation as zero difference between judges is the proper empirical expectation when judges are randomly assigned to cases.²⁴

²²One notable example provided in the literature is the placing of résumés on the “yes” or “no” pile (Bertrand & Mullainathan 2004).

²³It should be noted here that while the multilevel modeling in the analysis of detention/release was found to be substantively important, this is not the case for the analysis of detention length, suggesting that individual judges differ more with regard to the former decision.

²⁴It should be noted that another approach to uncovering judicial biases has been recently offered by Rachlinski et al. (2009), by directly testing for implicit racial bias and estimating its effect on simulated legal situations. The

We have statistically established that ethnic identity makes a difference in judicial decisions; however, one limitation of our study is that in the absence of a known ethnically-neutral baseline, we cannot formally determine whether our findings reflect discrimination against Arab suspects by Jewish judges, discrimination in favor of Arab suspects by Arab judges, or a mixture of both.

Still, in-group bias is not only a methodological bridge across the pitfalls of unobserved case characteristics. It is also theoretically important as it provides a mechanism for understanding discrimination against minorities (Greenwald & Krieger 2006). Since most judges belong to the ethnic and racial majority group in their jurisdiction, it is likely that the aggregate effect of in-group bias results in an overall discrimination against minorities in a national judicial system. In other words, systemic ethnic and racial disparities against minorities may simply result from many more majority judges practicing in-group bias compared to minority judges. Further research is required to establish the characteristics of in-group bias in other legal systems, in other legal procedures, the determinants of these biases, and their psychological mechanisms.

The existence of ethnic in-group bias in the early stages of the criminal process may also explain the difficulty in establishing the existence of ethnic discrimination in sentencing, since these early decisions, such as those regarding detention of suspects or defendants, can indirectly affect the entire process. Thus, if Arabs are more likely to be detained compared to Jews—as a result of translating individual-level in-group bias into aggregate-level discrimination against minorities—and if the detention influences the severity of the final sentence, then research on sentencing may attribute the increased severity of sentences for Arabs to their greater likelihood of prior detention rather than to their being Arabs. It is possible that this is the reason that studies that did not control for prior detention have found discrimination in sentencing among Arabs and Jews, while studies that controlled for it did not.²⁵ As noted earlier, the concentration on an early stage of the criminal process allows us to remove a considerable amount of “noise” from the analysis, and the potential to quantitatively control for the state’s position by relying on the police request enables us to address this important factor. To conclude, our findings robustly support the existence of judicial ethnic in-group bias in the decision to detain or release, yet not in the decision on the length of detention.

researchers administered the Implicit Association Test (IAT) to a sample of U.S. trial judges and then asked them to adjudicate a number of hypothetical cases. Their findings show that judges harbor implicit racial biases that are similar to those prevalent in the general U.S. population and that these biases can influence their judgment when they are unaware of a need to monitor their decisions for racial bias. Yet when judges are aware of this, they appear to be able to suppress that bias.

²⁵Compare Rattner and Fishman (1998) and Fishman et al. (2006), which found ethnic discrimination without controlling for prior detentions, to Hassin and Kremnitzer (1993), which found that prior detentions accounted for a considerable part of the variance in sentencing and did not find evidence for ethnic discrimination.

REFERENCES

- Abrams, David, Marianne Bertrand, & Sendhil Mullainathan (2008) *Do Judges Vary in Their Treatment of Race?* Mimeo. Chicago, IL: Univ. of Chicago.
- Albonetti, C. A. (1991) "An Integration of Theories to Explain Judicial Discretion," 38 *Social Problems* 247.
- Angrist, Joshua D., & Jorn-Steffen Pischke (2009) *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton/Oxford: Princeton Univ. Press.
- Ayres, I., & J. Waldfogel (1994) "A Market Test for Race Discrimination in Bail Setting," 46 *Stanford Law Rev.* 987.
- Bertrand, Marianne, Dolly Chugh, & Sendhil Mullainathan (2005) "Implicit Discrimination," 95(2) *American Economic Rev.* 94.
- Bertrand, Marianne, & Sendhil Mullainathan (2004) "Are Emily and Greg More Employable than Lakisha and Jamal?" 94(4) *American Economic Rev.* 991.
- Bridges, G. S., & S. Steen (1998) "Racial Disparities in Official Assessments of Juvenile Offenders: Attributional Stereotypes as Mediating Mechanisms," 63 *American Sociological Rev.* 554.
- Bushway, S. D., & A. M. Piehl (2001) "Judging Judicial Discretion: Legal Factors and Racial Discrimination in Sentencing," 35 *Law & Society Rev.* 733.
- Bynum, T. (1982) "Release on Recognizance: Substantive or Superficial Reform?" 20 *Criminology* 67.
- Chiricos, T. G., & W. D. Bales (1991) "Unemployment and Punishment: An Empirical Assessment," 29 *Criminology* 701.
- Chiricos, T. G., & C. Crawford (1995) "Race and Imprisonment: A Contextual Assessment of the Evidence," in *Ethnicity, Race, and Crime*. Albany, NY: State Univ. of New York Press.
- Clarke, S. H., & S. T. Kurtz (1983) "The Importance of Interim Decisions to Felony Trial Court Dispositions," 74 *J. of Criminal Law & Criminology* 475.
- Cohen, B. Z. (1985) "Judicial Discretion and Sentencing Disparity in Adult Felony Courts in Israel," 13 *J. of Criminal Justice* 99.
- Cohen, B. Z., & S. Palmor (1985) "The Impact of Ethnicity on Probation Recommendations and Dispositions in Israel," 2 *Justice Q.* 197.
- Daly, K., & M. Tonry (1997) "Gender, Race and Sentencing," 22 *Crime & Justice* 201.
- Davis, A. J. (1998) "Prosecution and Race: The Power and Privilege of Discretion," 67 *Fordham Law Rev.* 13.
- Demuth, S. (2003) "Racial and Ethnic Differences in Pretrial Release Decisions and Outcomes: A Comparison of Hispanic, Black, and White Felony Arrestees," 41 *Criminology* 873.
- Demuth, S., & D. Steffensmeier (2004a) "Ethnicity Effects on Sentence Outcomes in Large Urban Courts: Comparisons Among White, Black, and Hispanic Defendants," 85 *Social Science Q.* 994.
- (2004b) "The Impact of Gender and Race-Ethnicity in the Pretrial Release Process," 51(2) *Social Problems* 222.
- Englich, B., & T. Mussweiler (2001) "Sentencing Under Uncertainty: Anchoring Effects in the Courtroom," 31 *J. of Applied Social Psychology* 1535.
- Englich, B., T. Mussweiler, & F. Strack (2005) "The Last Word in Court—A Hidden Disadvantage for the Defense," 29 *Law & Human Behavior* 705.
- Fishman, G., A. Rattner, & H. Turjeman (2006) "Sentencing Outcomes in a Multinational Society: When Judges, Defendants and Victims Can Be Either Arabs or Jews," 3 *European J. of Criminology* 69.
- Fishman, G., A. Rattner, & G. Weimann (1987) "The Effect of Ethnicity on Crime Attribution," 25 *Criminology* 507.
- Frazier C. E., & D. M. Bishop (1985) "The Pretrial Detention of Juveniles and Its Impact on Case Dispositions," 76 *J. of Criminal Law & Criminology* 1132.
- Green, William H. (1994) *Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models*, NYU Working Paper No. EC-94-10. Available at <<http://ssrn.com/abstract=1293115>>.
- Greenwald, Anthony G., & Linda Hamilton Krieger (2006) "Implicit Bias: Scientific Foundations," 94(4) *California Law Rev.* 945.

- Haj Yahia, M. S., G. Rahav, & M. Teichman (1994) "The Lower Court and Minorities in Israel," 4 *Plilim: Israel J. of Criminal Justice* 157 (in Hebrew).
- Hassin, Y., & M. Kremnitzer (1988) "Uniformity in Adjudicating Traffic Violators," 18(1) *Mishpatim* 103 (in Hebrew).
- (1993) "The Offence of Breaking and Entering—Variables Affecting Consistency in Sentencing," 22 *Mishpatim* 571 (in Hebrew).
- Heckman, James J. (1979) "Sample Selection Bias as a Specification Error," 45(1) *Econometrica* 153.
- Holmes, M. D., H. M. Hosch, H. C. Daudistel, D. A. Perez, & J. B. Graves (1996) "Ethnicity, Legal Resources, and Felony Dispositions in Two Southwestern Jurisdictions," 13 *Justice Q.* 11.
- Humphrey, J., & T. J. Fogarty (1987) "Race and Plea Bargained Outcomes: A Research Note," 66 *Social Forces* 176.
- Johnson, Brian D. (2006) "The Multilevel Context of Criminal Sentencing: Integrating Judge- and County-Level Influences," 44(2) *Criminology* 259.
- Jolls, Christine, & Cass R. Sunstein (2006) "The Law of Implicit Bias," 94(4) *California Law Rev.* 969.
- Kahneman, Daniel, & Shane Frederick (2002) "Representativeness Revisited: Attribute Substitution in Intuitive Judgment," in T. Gilovich, D. Griffin, & D. Kahneman, eds., *Heuristics and Biases*, pp. 49–81. New York: Cambridge Univ. Press.
- King, Gary (1991) "Truth is Stranger than Prediction, More Questionable than Causal Inference," 35(4) *American J. of Political Science* 1047.
- Kleck, G. (1981) "Racial Discrimination in Criminal Sentencing: A Critical Evaluation of the Evidence with Additional Evidence on the Death Penalty," 46 *American Sociological Rev.* 783.
- Klein, Stephen, Joan Petersilia, & Susan Turner (1990) "Race and Imprisonment Decisions in California," 247(4944) *Science* 812.
- Kretzmer, D. (1990) *The Legal Status of Arabs in Israel*. Boulder, CO: Westview Press.
- Patterson, E., & M. Lynch (1991) "Bias in Formalized Bail Procedures," in M. Lynch & E. Patterson, eds., *Race and Criminal Justice*, pp. 36–53. Albany, NY: Harrow & Heston.
- Petee, T. (1994) "Recommended for Release on Recognizance: Factors Affecting Pretrial Release Recommendations," 134 *J. of Social Psychology* 375.
- Puhani, Patrick A. (2000) "The Heckman Correction for Sample Selection and its Critique," 14(1) *J. of Economic Surveys* 53.
- Rachlinski, Jeffrey J., Sheri Lynn Johnson, Andrew J. Wistrich, & Chris Guthrie (2009) "Does Unconscious Racial Bias Affect Trial Judges?" 84(3) *Notre Dame Law Rev.* 1195.
- Rankin, A. (1964) "The Effect of Pretrial Detention," 39 *New York Univ. Law Rev.* 641.
- Rattner, Arye, & Gideon Fishman (1998) *Justice for All? Jews and Arabs in the Israeli Criminal Justice System*. Westport, CT: Praeger.
- Samoocha, Sammy (2005) *Index of Arab-Jewish Relations in Israel 2004*. Haifa/Jerusalem/Tel Aviv: Jewish-Arab Center, Univ. of Haifa/Citizens' Accord Forum Between Jews and Arabs in Israel/Friedrich Ebert Stiftung.
- Sampson, R. J., & J. L. Lauritsen (1997) "Racial and Ethnic Disparities in Crime and Criminal Justice in the United States," 21 *Crime & Justice* 311.
- Schlesinger, T. (2005) "Racial and Ethnic Disparity in Pretrial Criminal Processing," 22 *Justice Q.* 192.
- Shadish, William R., Thomas D. Cook, & Donald T. Campbell (2002) *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston/New York: Houghton Mifflin Company.
- Spohn, C. C. (1990a) "Decision Making in Sexual Assault Cases: Do Black and Female Judges Make a Difference?" 2 *Women & Criminal Justice* 83.
- (1990b) "The Sentencing Decisions of Black and White Judges: Expected and Unexpected Similarities," 24 *Law & Society Rev.* 1197.
- (2000) "Thirty Years of Sentencing Reform: The Quest for a Racially Neutral Sentencing Process," in J. Horney, ed., *Criminal Justice 2000, Vol. 3. Policies, Processes, and Decisions of the Criminal Justice System* (pp. 427–501). Washington, DC: U.S. Department of Justice, National Institute of Justice.
- Steenbergen, Marco R., & Bradford S. Jones (2002) "Modeling Multilevel Data Structures," 46(1) *American J. of Political Science* 218.

Steffensmeier, D., & C. L. Britt (2001) "Judges' Race and Judicial Decision-Making: Do Black Judges Sentence Differently?" 82 *Social Science Q.* 749.

Tonry, Michael H., ed. (1995) *Malign Neglect—Race, Crime, and Punishment in America*. New York: Oxford Univ. Press.

——, ed. (1997) *Ethnicity, Crime, and Immigration: Comparative and Cross-National Perspectives*. Chicago, IL: Univ. of Chicago Press.

Uhlman, T. M. (1978) "Black Elite Decision Making: The Case of Trial Judges," 22 *American J. of Political Science* 884.

Warner, D. E. (2000) "Race and Ethnic Bias in Sentencing Decisions: A Review and Critique of the Literature," in Michael W. Markowitz & Delores D. Jones-Brown, eds., *The System in Black and White: Exploring the Connection Between Race, Crime and Justice*. Westport, CT: Praeger.

Welch, S., M. Combs, & J. Gruhl (1988) "Do Black Judges Make a Difference?" 32 *American J. of Political Science* 126.

Williams, M. R. (2003) "The Effect of Pretrial Detention on Imprisonment Decisions," 28 *Criminal Justice Rev.* 299.

Zussman, A., & M. Shayo (2010) Judicial Ingroup Bias in the Shadow of Terrorism, unpublished manuscript, available at <<http://pluto.huji.ac.il/~azussman.couts.pdf>>.

APPENDIX A: FURTHER RANDOMIZATION TESTS

Dependent variable: logit(Arab judge).
 Each cell reports *p* values from Wald test.

	1	2	3	4	5
Arab suspect	0.579	0.604	0.324	0.494	0.481
Police request (in days)	0.990	0.931	0.854	0.951	0.832
Mid-low severity		0.554	0.129	0.126	0.136
Mid-high severity		0.785	0.499	0.441	0.463
High severity		0.674	0.266	0.230	0.223
Violent offenses			0.757	0.741	0.727
Illegal residence			0.077	0.093	0.079
Sexual offenses			0.981	0.938	0.889
Domestic violence			0.217	0.206	0.216
Drug-related offenses			0.549	0.503	0.522
Traffic violations			0.950	0.941	0.815
Public defender				0.139	0.176
Unrepresented				0.475	0.557
Prosecutor's statement					0.162
Day (Saturday)					0.966
Constant	<0.001	<0.001	<0.001	0.002	0.002
χ^2	0.31	0.67	6.88	9.05	10.98
−2 Log likelihood	1,311.894	1,311.529	1,304.810	1,302.636	1,300.706
Pseudo <i>R</i> ²	0.000	0.000	0.005	0.007	0.008

N = 1,230.

APPENDIX B: ESTIMATING THE EFFECTS OF ETHNIC IN-GROUP ON DETENTION LENGTH—LINEAR AND MULTILEVEL LINEAR REGRESSIONS

	9	10	11	12
	b	b	b	b
Arab judge × Arab suspect	0.223 (0.206)	0.166 (0.161)	0.192 (0.159)	0.117 (0.156)
Arab judge	−0.080 (0.139)	−0.103 (0.111)	−0.113 (0.110)	−0.094 (0.159)
Arab suspect	0.021 (0.101)	−0.066 (0.081)	0.056 (0.084)	0.009 (0.081)
Judge gender (male)		0.317*** (0.073)	0.276*** (0.073)	0.208* (0.121)
Police’s request		0.424*** (0.020)	0.419*** (0.020)	0.417*** (0.019)
Court (Nazareth)		0.019 (0.083)	−0.013 (0.084)	−0.018 (0.134)
Court (Acre)		−0.064 (0.088)	−0.143 (0.089)	−0.100 (0.164)
Mid-low severity		−0.148 (0.091)	−0.057 (0.110)	−0.019 (0.105)
Mid-high severity		0.090 (0.125)	0.047 (0.132)	0.014 (0.126)
High severity		0.050 (0.111)	−0.195 (0.137)	−0.174 (0.131)
Violent offense			0.164 (0.100)	0.213** (0.096)
Illegal residence			0.135 (0.204)	0.103 (0.196)
Sexual offense			0.308 (0.193)	0.320* (0.184)
Domestic violence			0.024 (0.105)	−0.006 (0.100)
Drug-related offense			0.492*** (0.153)	0.489** (0.148)
Traffic violation			−0.099 (0.300)	−0.020 (0.286)
Public defender			−0.065 (0.083)	−0.020 (0.080)
Unrepresented			0.016 (0.133)	0.040 (0.130)
Prosecutor’s declaration			−0.018 (0.128)	0.010 (0.123)
Day (Saturday)			−0.356*** (0.073)	−0.292*** (0.074)
Inverse Mills ratio				
Constant	2.836*** (0.069)	0.520*** (0.134)	0.661*** (0.172)	0.574*** (0.190)
Likelihood-ratio test of sigma				30.64***
N	896	896	896	896

*, **, and *** denote significance levels of 10 percent, 5 percent, and 1 percent, respectively.

APPENDIX C: ESTIMATING THE EFFECT OF ETHNIC IN-GROUP
ON DETENTION LENGTH—ROBUSTNESS TESTS

	<i>Zero-Inflated Poisson</i>	<i>Heckman Two-Step Selection Model—Linear Model</i>	<i>Multilevel Poisson with Heckman Selection Correction</i>
Detention length			
Arab judge × Arab suspect	−0.011 (0.097)	0.173 (0.157)	0.075 (0.093)
Arab judge	0.008 (0.068)	−0.095 (0.108)	−0.047 (0.064)
Arab suspect	0.035 (0.045)	−0.060 (0.079)	−0.016 (0.047)
Judge gender (male)	0.071 (0.045)	0.265*** (0.072)	0.096** (0.043)
Police’s request	0.144*** (0.009)	0.418*** (0.019)	0.119*** (0.009)
Violent offense	0.042 (0.055)	0.139 (0.090)	0.049 (0.053)
Illegal residence	−0.009 (0.126)	0.145 (0.182)	0.014 (0.119)
Sexual offense	0.071 (0.109)	0.266 (0.182)	0.099 (0.102)
Domestic violence	−0.048 (0.067)	−0.006 (0.099)	0.002 (0.062)
Drug-related offense	0.134** (0.067)	0.366*** (0.116)	0.130* (0.067)
Traffic violation	−0.363* (0.189)	−0.104 (0.285)	−0.031 (0.178)
Day (Saturday)	−0.168*** (0.045)	−0.355*** (0.071)	−0.145*** (0.043)
Inverse Mills ratio			0.064 (0.144)
Constant	0.119 (0.076)	0.521*** (0.183)	0.332*** (0.107)
Release			
Court (Nazareth)	−2.66*** (0.937)	0.531*** (0.104)	
Court (Acre)	−0.345 (0.220)	0.148 (0.092)	
Mid-low severity	−0.341 (0.234)	0.179* (0.094)	
Mid-high severity	−1.502*** (0.527)	0.627*** (0.154)	
High severity	−1.294*** (0.380)	0.616*** (0.128)	
Public defender	−0.681*** (0.235)	0.095 (0.093)	
Unrepresented	−0.491 (0.356)	0.081 (0.149)	
Prosecutor’s declaration	−14.191 (669.76)	0.245 (0.166)	
Constant	−0.088 (0.266)	0.105 (0.115)	
<i>N</i>	1,229	1,229	896

*, **, and *** denote significance levels of 10 percent, 5 percent, and 1 percent, respectively.